

# Updated Methodologies for Estimating Fuel Consumption Using the 2017 National Household Travel Survey

Estimates of fuel economy, fuel consumption, and fuel costs in the 2017 National Household Travel Survey (NHTS) were calculated by the Energy Information Administration (EIA) largely using the same process as in the 2009 NHTS. The changes for 2017 listed below serve as amendments to the 2009 documentation (available at <https://nhts.ornl.gov/2009/pub/DerivedAddedVariables2009.pdf>) and explain changes in methodology for future use.

## User-Supplied Make and Model Names

In prior surveys, the NHTS collected make and model names from users during the telephone interview, after which the interviewer selected pre-determined categories if they were available. In all cases, the original response was retained and available for use. However, the 2017 NHTS was conducted both online and via telephone, and only pre-determined make and model categories were collected. This provided less granular information than in past surveys for use in the matching process against outside fuel economy databases.

## Corporate Average Fuel Economy (CAFE)

Commonly referred to as CAFE, the Corporate Average Fuel Economy, was created by the National Highway Traffic Safety Administration (NHTSA) in conjunction with the Environmental Protection Agency (EPA). CAFE has been used in years past to create the EIA variables that are attached to the NHTS. CAFE data contains a sales-weighted average fuel economy, expressed in miles per gallon. Since the CAFE estimate is, in part, a regulatory tool that “reflects various credits, incentives, and adjustments available to automakers<sup>1</sup>,” this estimate was then adjusted in 2001 and 2009 NHTS surveys to reflect real-world use. For 2017, however, EIA employed a different method due to the CAFE being unavailable for use. EPA’s fueleconomy.gov data were used in CAFE’s place.

## Fueleconomy.gov

Fueleconomy.gov has a list of all makes and models of vehicles back to 1984. These vehicles have minimum, maximum and average fuel economy values for city driving, highway driving and combined driving conditions. Due to the unavailability of the CAFE data for the development of the 2017 NHTS variables, fueleconomy.gov was substituted. From the vehicles that would be matched from the 1-to-25 threshold (see *Matching Make/Model Between NHTS and Fueleconomy.gov* below), these vehicles would then be averaged together to create minimum fuel consumption, maximum fuel consumption and average fuel consumption, with average fuel consumption being the measure included in the NHTS data. Note that unlike CAFE, fueleconomy.gov estimates did not need adjustment to reflect real-world

---

<sup>1</sup> <https://nepis.epa.gov/Exe/ZyPDF.cgi/P100IENA.PDF?Dockey=P100IENA.PDF>, accessed May 15, 2020.

use, as they are already calculated to reflect real-world use. Of note, the estimates found in the 2017 NHTS reflect the 2008 fueleconomy.gov methodology.<sup>2</sup>

## **Matching Make/Model Between NHTS and Fueleconomy.gov**

The matching process refers to taking NHTS make and model information and matching it with that found in an outside database, such as CAFE in 2001 and 2009 or fueleconomy.gov in 2017. Due to full make and model names not being explicitly collected in the 2017 NHTS, the matching process was far more challenging. While using standardized make and model categories has some advantages, in the case of matching vehicles from fueleconomy.gov to the NHTS made things more difficult when information was withheld or missing. Matching would range from perfect matches 1-to-1 to not so perfect matches of 1-to-25. 1-to-1 matches were matched perfectly from fueleconomy.gov makes and models and NHTS sampled vehicles. If more than one vehicle is a possible match from fueleconomy.gov, the min would be the most efficient possible match, the max would be least efficient of the possible matches, and the average is the arithmetic average of all the possible matches. As there were no sales data available, there was no weighting of the possible matches to calculate the arithmetic average. This average is what is included in the NHTS vehicle file.

## **Hot-Decking/Imputation**

Survey respondents may provide all, some, or no information about their vehicle. When information is missing, it becomes harder to match vehicles to the appropriate fuel economy data. To aid in finding the appropriate match, a multi-step imputation approach was used. First, vehicles with a missing fuel (FUELTYPE) and vehicle type (VEHTYPE) were imputed. If a vehicle was imputed to a hybrid, then the appropriate hybrid fuel (HFUEL) was imputed as well. Next, fuel economy data was imputed using hot-deck imputation. A vehicle with missing fuel economy data is called a recipient, and a vehicle with all non-missing fuel economy that provides the values for imputation is called a donor. Similar vehicles were grouped together into classes by certain characteristics based on their fuel type. Regression models were used to determine the best class variables for each fuel. Data was imputed if no matches or too many matches were found for a vehicle. If a vehicle had too many matches, this was an indication that the matching process<sup>3</sup> could not find an ideal match. As a result, the fuel economy of any vehicle with more than 25 matches was imputed as well.

## **Motorcycles, Motor Homes, and Heavy Trucks**

In previous years there were imputations for motorcycles, motor homes and heavy trucks because sales data was provided for these vehicles. Because of the switch to fueleconomy.gov data and the lack of sales data, these vehicles were left out of the consumption calculations as they were not represented in the fueleconomy.gov data.

---

<sup>2</sup> An update to fueleconomy.gov's methodology was also performed in 2017 (see <https://fueleconomy.gov/feg/ratings.shtml>, accessed May 15, 2020.) Since this new 2017 methodology was only applied to vehicles with model years of 2011 and later, it was not used in calculating the 2017 NHTS estimates.

<sup>3</sup> Performed in Excel using Visual Basic for Applications (VBA)

## Compatibility with Prior Surveys

Because of changes in methodology used to create the 2017 NHTS consumption metrics, compatibility with prior surveys should not be assumed. The fueleconomy.gov numbers are designed to give consumers real-world estimates, and thus *could be* reasonably comparable, without adjustment, to the adjusted CAFE estimates in 2001 and 2009. However, caution is still advised in interpreting any differences between 2009 and 2017, since the source databases are different.